DOCUMENT RESUME

ED 305 597 CS 009 587

AUTHOR Lysynchuk, Linda; And Others
TITLE Reciprocal Instruction Improves Standardized Reading Comprehension Performance in Poor Grade-School Comprehenders.
PUB DATE 27 Jan 89
NOTE 41p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, March 27-31, 1989).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Foreign Countries; Grade 4; Grade 7; Intermediate Grades; Junior High Schools; Reader Text Relationship; *Reading Comprehension; Reading Difficulties; Reading Improvement; *Reading Instruction; Reading Research; *Reading Strategies
IDENTIFIERS Canada; *Reciprocal Teaching

ABSTRACT

A study hypothesized that if prediction, clarification, question-generation, and summarization strategies were taught using scaffolded instruction, then poor comprehenders might be able to learn the strategies so that they could apply them to text on their own and thereby improve their performance on standardized comprehension measures. Subjects were 36 fourth grade and 36 seventh grade poor comprehenders (all were English-speaking Canadians) who participated in 13 sessions of reading strategy instruction or reading practice (control condition). Students were instructed to make predictions when reading, to generate questions about text, to summarize what was read, and to clarify points that were hard to understand. The strategies were first taught using the reciprocal instruction approach developed by A. S. Palincsar and A. L. Brown, which involves providing support to students as they need it and peer teaching of strategies. Control subjects were exposed to the same materials as reciprocally trained students, but were given no strategy instructions. The most important finding was that a greater increase from before to after training on standardized reading comprehension occurred in the subjects who were reciprocally trained than in the control groups. This effect is important, given that standardized comprehension measures tap important reading skills and are used extensively in making curriculum and instruction decisions. Additional research in reciprocal instruction would be useful. (Five tables of data and one figure are included, and 43 references are attached.) (RAE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED305597

Reciprocal Instruction Improves Standardized Reading Comprehension Performance in Poor Grade-School Comprehenders

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Linda Lysynchuk

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Linda Lysynchuk Michael Pressley

University of Western Ontario

Nancy J. Vye

Vanderbilt University

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

Running Head: Reciprocal Instruction

This research was funded by an Ontario Graduate Scholarship to the first author and a grant to the second author from the Natural Sciences and Engineering Research Council of Canada. Thanks to Jacqueline A. McGoldrick and Patrick Rocco who along with the first author served as instructors in the study. The London, Ontario Board of Education and Dr. Richard Stennett of the Board provided outstanding cooperation during the conduct of the study. Vicky Martin, Sonya Symons, Teena Willoughby, Vera Woloshyn, and Eileen Wood provided invaluable input on previous versions of this article. Correspondence about the article should be addressed to Michael Pressley, Department of Psychology, University of Western Ontario, London, Ontario N6A 5C2 Canada.

A preliminary shorter version of this study that was circulated for discussion and comments contained errors. This version of 27 January 1989 replaces all previous versions, which we request that you destroy.

CS009587

BEST COPY AVAILABLE

Abstract

Grade-4 and -7 poor comprehenders participated in 13 sessions of reading-strategy instruction or reading practice (control condition). Trained students were instructed to make predictions when reading, to generate questions about text, to summarize what was read, and to clarify points that are hard to understand. The strategies were taught using the reciprocal instruction approach developed by Palincsar and Brown (1984), which involves providing support to students as they need it and peer teaching of strategies. Control subjects were exposed to the same materials as reciprocally-trained students, but were given no strategy instructions. The most important finding was a greater increase from before to after training on standardized reading comprehension in the reciprocally-trained than in the control condition. This effect is important given that standardized comprehension measures tap important reading skills and are used extensively in making curriculum and instruction decisions. The importance of additional research on reciprocal instruction or revised versions of it is considered briefly.

Reciprocal Instruction Improves Standardized Reading Comprehension Performance in Poor Grade-School Comprehenders

One of the most notable strategy instruction packages developed in the 1980s (e. g., Pearson & Dole, 1987; Pressley, Snyder, & Cariglia-Bull, 1987) was reciprocal instruction (Palincsar & Brown, 1984), a method of reading instruction designed to improve comprehension in children who can decode but who experience difficulty understanding text. The intervention involves training of four strategies that are associated with both skilled comprehension and comprehension monitoring while reading (e. g., August, Flavell, & Clift, 1984; Bereiter & Bird, 1985; Meyer, Brandt, & Bluth, 1980). Reciprocal instruction encourages student self-directed prediction of what information might occur in text, clarification of information not completely understood as text is read, generation of questions about text content, and summarization of material covered.

How does articulation of these strategies promote understanding and awareness of comprehension? Making predictions should activate prior knowledge and create expectations, increasing the meaningfulness and memorability of the text (e. g., Anderson & Pearson, 1984); seeking clarifications when unsure of what content means promotes both monitoring of comprehension difficulties and use of reprocessing strategies like selective search of text for relevant content and rereading (e. g., Baker, 1985); generating questions promotes integration of text (Davey & McBride, 1986); and summarizing promotes analysis and selective encoding of text (e. g., Brown & Day, 1983). Attempting to summarize can also make obvious whether material has been understood completely (and thus, can be a strong cue about whether there is a need to reprocess),

since it should be difficult to construct a summary if text has not been understood and remembered (Rinehart, Stahl, & Erickson, 1986).

A hallmark of reciprocal instruction is that teaching is scaffolded (e. g., Manzo, 1969; Wood, Bruner, & Ross, 1976), particularly along the lines suggested by Vygotsky (1978). Vygotsky argued that much of cognitive development is mediated by experts providing input to novices (e. g., parents to their children, teachers to their students), with the provision of expert support to children gradually diminishing as pupil competence increases (i. e., just as scaffolding is gradually removed from a building as it is progressively better able to support its own weight). First, the expert completely guides the child's activity, modeling how the task can be performed. As an observer during this initial phase, the child actually does very little cognitive work. The child then attempts the task, with the adult providing supportive cuing, assistance, and additional modeling as required. Eventually the child assumes most of the thinking responsibilities. The adult is now a supportive and sympathetic audience. A case can be made that some comprehension-fostering activities are already passed from experts to children using scaffolded instruction as described by Vygotsky (1978). For instance, when parents interact with children as they "read" picture books, the adult initially directs the comprehension process, with the procedures modeled by the adult eventually internalized by the child (e. g., Ninio & Bruner, 1978). Scaffolded instruction of more mature comprehension strategies is not a regular part of reading instruction, however (e. g., Durkin, 1979).

Palincsar and Brown (1984) devised a way to scaffold such teaching, at least for students who are already experiencing comprehension difficulties. They described scaffolded instruction of the four

comprehension strategies described earlier, in this case when teaching a single student:

> ...[the] adult teacher...assigned a segment of [a] passage to be read and either indicated that it was her turn to be the teacher or assigned the student to teach the segment. The adult teacher and the student then read the assigned segment silently. After reading the text, the teacher (student or adult) for that segment asked a question that a teacher or a test might ask on the segment, summarized the content, discussed and clarified any difficulties, and finally made a prediction about future content. All of these activities were embedded in as natural a dialogue as possible, with the teacher and student giving feedback to each other.
>
> Initially, the adult teacher modeled the activities, but the students had great difficulty assuming the role of dialogue leader when their turn came. The adult teacher was sometimes forced to construct paraphrases and questions for the students to mimic. In this initial phase, the adult teacher was modeling effective comprehension-monitoring strategies, but the students were relatively passive observers.
>
> Gradually, the students became much more capable of assuming their role as dialogue leader and by the end of ten sessions were providing paraphrases and questions of some sophistication. (pp. 124-125)

In short, teaching proceeded from teacher-controlled to student-

controlled, with modeling, explanations, and feedback provided to students as needed.

In addition to developing reciprocal instruction, Palincsar and Brown (1984) also provided the most extensive evaluation of it. All of their research involved junior-high-school students who were experiencing comprehension problems. Training was conducted for 20 days, with pairs of students (Experiment 1) and whole classes (Experiment 2). Although the results were somewhat complicated, in general, Palincsar and Brown (1984) demonstrated that reciprocal training did in fact improve pretest-to-posttest comprehension in this population. Without a doubt, one of the most striking improvements was an increase from before training to after training on a standardized reading comprehension measure. This finding in particular was received skeptically by some. For instance, consider Carver's (1987) somewhat sarcastic reaction:

> ...the data collected by Palincsar and Brown [Experiment 1] seem to suggest that we can administer the reciprocal teaching procedure to seventh-grade children who are reading at fourth-grade level (with a reported mean IQ of 83) and they will be equal in reading ability to average seventh graders in 20 days (requiring only 30 minutes a day). This finding was described by Pearson [1985] as "dramatic"; even that description seems modest. I would describe my first reaction to this treatment effect as breathtaking because of the incredibly large size. (p. 120).

Although our reading of the evidence was that Carver erred on the side of hyperbole (e. g., standardized reading achievement improved by 3 years for only 1 of the 6 students, with an average achievement gain of

1.25 years), the overall improvement in comprehension from before to after reciprocal instruction was striking in Palincsar and Brown's (1984) Experiment 1. We believed that such a promising procedure deserved study by other researchers than its developers. Thus, the purpose of the study reported here was to provide an evaluation of reciprocal instruction in a true experiment.

The main hypothesis in this experiment was that if prediction, clarification, question-generation, and summarization strategies were taught using scaffolded instruction (i. e., if Palincsar & Brown's, 1984, reciprocal instruction package were provided), poor comprehenders might be able to learn the strategies so that they could apply them to text on their own . If so, performance on standardized comprehension measures should increase.

We focussed on standardized comprehension in this investigation because it was the only comprehension measure in Palincsar and Brown (1984) that is widely used in schools. Not only are such measures familiar to educators, they are often important indices of accountability, and thus, information about reciprocal instruction effects on standardized test performance could be especially helpful to those charged with making curriculum and instruction decisions. Finally, it is known that standardized comprehension performance is not easily improved (e. g., Haller, Child, & Walberg, 1988). When an intervention produces true gains on standardized comprehension, it is an impressive demonstration of an intervention's potency.

A main strength of this study is that there was better collection of standardized reading comprehension data compared to Palincsar and Brown's (1984) Experiment 1. First, the pre-intervention standardized

reading data in the previous investigation were generated by school officials using group testing procedures. In contrast, the post-intervention data were collected on a one-to-one basis by a researcher. Thus, it is difficult to know if all of the pre- to posttest performance improvement in standardized comprehension was due to reciprocal instruction. Some of it might have been due to the greater sensitivity of one-to-one assessment compared to group-administered testing. Second, none of the subjects in non-reciprocal conditions in Palincsar and Brown's (1984) study were administered the standardized reading measure following their participation in the study. Thus, only an uncontrolled pretest-to-posttest evaluation could be made of the reading comprehension achievement data in the reciprocally-trained condition (Campbell & Stanley, 1966). Third, the post-intervention standardized reading scores were collected three months after the intervention had ended, so that some of the gains in standardized comprehension could have been due to schooling effects subsequent to treatment. In contrast, we used the same standardized testing procedures both before and after the intervention: All participants in the present study were tested in the same fashion, with the post-treatment evaluation made immediately after the last treatment session.

The basic design of the experiment reported here was very simple. There was one reciprocal instruction and one control group. The control participants were treated as much like the reciprocal-instruction group as possible, except that they were given no reciprocal teaching of comprehension strategies. That is, controls met the same number of times with the same instructor who saw the reciprocally-trained students; they were exposed to the same practice materials as reciprocally-trained

subjects; and they took the same practice tests.

This control condition differs from the one used in Palincsar and Brown (1984, Experiment 1). In that study, control students were seen only on occasions corresponding to the beginning and end of reciprocal training. They were administered the same pretest and posttest assessments as reciprocally-trained students, but otherwise did not receive contact with the experimenters. Why was a maximally-similar control condition used here rather than the no-instruction control employed by Palincsar and Brown? Comparison of a reciprocally-trained condition with a maximally-similar control condition permits the conclusion that any trained versus control condition differences are due to the reciprocal instruction per se. In contrast, trained versus no-instruction control differences could be due to social demand factors produced by differential contact with the experimenter or to greater exposure to training materials in the reciprocal instruction condition.

Palincsar and Brown (1984, Experiment 1) reported both standardized measures of comprehension and vocabulary at the beginning and the end of the study, in an effort to establish the discriminative effectiveness of reciprocal instruction. Thus, we included both comprehension and vocabulary measures here. Reciprocal teaching was hypothesized to affect comprehension processes and thus, pretest-to-posttest improvements for standardized comprehension were expected to be greater in the reciprocally-trained condition than in the control group. On the other hand, there was no theoretical reason to expect that comprehension instruction would have a dramatic or differential effect on standardized vocabulary performance. Even greatly improved comprehension for the several weeks of reciprocal instruction would be expected to have little

effect on long-term knowledge of vocabulary. Thus, it was anticipated that there would only be small pretest-to- posttest changes in both conditions for standardized vocabulary achievement. If this pattern were obtained, it would be strong evidence that reciprocal instruction caused the standardized comprehension gains. Alternative causal mechanisms (e. g., increased motivation to do well on tests, improvements in test-taking skills) would be expected to have general rather than discriminative effects---That is, alternative causal mechanisms would be expected to affect both comprehension and vocabulary performances. See Cole and Means (1981) or Rohwer (1976) for additional commentary about isolation of causal mechanisms via analyses of patterns of discriminative effects.

Methods

Subjects

Thirty-six grade-4 (10 females, 26 males; age range = 9 yrs 2 mos to 10 yrs 10 mos) and 36 grade-7 students (22 females, 14 males; age range = 12 yrs 5 mos to 14 yrs 1 mo) participated in this study. The grade-4 students were enrolled in 6 schools; the grade-7 participants were in 2 schools. All were English-speaking Canadians. (Grade-4 and grade-7 students were studied here because comprehension is emphasized as an important goal of reading instruction throughout the second half of the grade-school years.)

As in Palincsar and Brown's (1984) Experiment 1, all participants were nominated by their teachers as adequate decoders but poor comprehenders. None were classified as learning disabled or mentally retarded, again consistent with Palincsar and Brown (1984, Experiment 1). Whether the subjects actually met the criteria of being adequate decoders but poor comprehenders was checked using standardized measures. All

could decode at least 80% of the words that students at their grade level are expected to be able to decode as measured by performance on the Diagnostic Reading Scales (Spache, 1972). All grade-4 subjects scored below the 50th percentile on the reading comprehension subtest of the Metropolitan Achievement Test (Prescott, Balow, Hogan, & Farr, 1978); all grade-7 participants scored below the 50th percentile on the reading comprehension subtest of the Gates- MacGinitie Reading Test (MacGinitie, 1978). Following the pretest, subjects at each grade level were paired on the basis of their pretest scores (two lowest children in a pair, next two lowest, and so on) with one pairmate randomly assigned to either the control or the reciprocal- training condition. Thus, at the beginning of the study, participants in the trained and control conditions were comparable with respect to standardized comprehension performance.

Materials and Measurements

Standardized assessments before and after the experiment. The Gates-MacGinitie was used in this experiment at the grade-7 level to be consistent with Palincsar and Brown (1984). Since there is no grade-4 version of that test, the Metropolitan Achievement Test was used at the lower grade level. The comprehension portions of these tests are similar in format. The tests include a number of short passages followed by several multiple-choice questions. The questions tap understanding of details as well as whether students can integrate information in the passages. Understanding of inferences and implications suggested by text are also covered by some items.

The vocabulary subtests of the Gates-MacGinitie Tests and Canadian Tests of Basic Skills (King, 1967) were also administered at the grade-7 and grade-4 levels respectively. Children are presented individual words

on these tests and are required to select synonyms given multiple choices.

All participants in a school were administered these standardized tests simultaneously (i. e., the tests were group administered), with all participants tested twice (immediately before and at the end of the experiment). Alternative forms of the tests were administered to each subject at the two testings, with appropriate counterbalancing. That is, half the grade-7 subjects received one form of the Gates-MacGinitie before the 13 experimental sessions, and half received the other form before the 13 sessions. Subjects were tested at the end of the experiment on the version of the test not experienced previously. The standardized comprehension measures required 45 mins per administration; the vocabulary assessment required 15 mins.

During the experiment. Grade-appropriate expository passages (300 to 900 words long) were used for training in the reciprocal instruction condition; control subjects read these passages during sessions corresponding to when reciprocal training occurred. The 13 grade-4 passages were taken from Bailey and Quick (1961), Code (1971), and Thorn and Richmond (1972); the 13 grade-7 passages were from Code (1971). In addition, the same sources provided 26 shorter (about 200 word) expository passages (i. e., one for each of 13 days of training at the two grade levels) that were used to make daily assessments. Evaluation of student comprehension (i. e., 10 comprehension questions were answered after a passage was read) alternated with assessments of whether subjects could retell the contents of passages from memory, so that half the daily assessments involved question answering and half retelling. The same passage was used in both conditions for a given training session. Two hundred word passages were used for daily assessments because of

time constraints. On those days when the daily assessment involved answering comprehension questions, 5 of the 10 questions required recall of something explicitly stated in the passage and 5 required report of something implicit in the text. See Appendix A for one such passage and its 10 comprehension questions.

In closing this subsection, we note that the training and daily assessment passages used in this study were shorter than the passages used by Palincsar and Brown (1984, Experiment 1). The length of passages was constrained by the amount of time available for each of the 13 sessions. We selected readings that would permit completion of the training passage and daily assessment in a half hour with a relaxed pace.

Summary. The various assessments and when they were made is summarized in Table 1.

Insert Table 1 About Here

Procedures

Participants were seen by one of three experimenters (two females, one male, all in their 20s). A given child worked consistently with only one of the experimenters. Each experimenter worked with an equal number of reciprocally-trained and control participants.

Reciprocally-trained condition. Strategy training occurred in small groups of 2 to 5 students over the course of 13 days. Although this was shorter than the 20 days of training in Palincsar and Brown (1984, Experiment 1), improvements in processing due to reciprocal instruction reached asymptote by the end of the thirteenth session in their study. Thus, we were optimistic that intervention gains might be obtained with

less than the full 20 days used in the previous study. The sequence and content of instruction was comparable to that used by Palincsar and Brown (1984, Experiment 1): (a) The teacher first explained and modeled coordinated use of the four comprehension strategies. (b) Students executed the strategies in a reading group, taking turns acting as the leader of the group. These student leaders elicited predictions, questions about the text, requests for clarifications, and summaries of text content from the reading group participants. (c) The adult teacher provided guidance and feedback as required. (d) There was also support from fellow students who provided input to each other as required. (e) Instruction was withdrawn as students' competence in executing strategies increased. (f) Throughout the process, the adult provided critical metacognitive information about the strategies, such as commentary about why, when, and where to use the procedures that were being trained.

The 13 sessions of reciprocal training were divisible into four phases: (a) On the first four days, it was explained to subjects how to ask questions, clarify unknown words or phrases, summarize passages, and predict what would happen next using passage clues. These explanations followed scripted lessons developed by Palincsar and Brown. Each day a new strategy was introduced and the previously learned strategies were practiced. Students took turns attempting the trained processes beginning with the first session. Rather than taking up large amounts of space relating the scripted commentary, readers are referred to Palincsar and Brown (1984), Palincsar (1986, 1987), and Palincsar, Brown, and Martin (1987) for detailed illustration of such explanations. The most important attributes of the four trained strategies are reviewed in Table 2, however.

Insert Table 2 About Here

(b) For days 5 and 6 subjects read the story aloud and took turns clarifying, questioning, summarizing, and predicting. On a typical day, the adult instructor introduced the training passage for the day with a brief discussion intended to activate students' prior knowledge about the topic. The adult called attention to the title and requested predictions about the passage based on it. Then, the adult assigned a student to serve as "teacher" for a segment of the passage. This student leader attempted to predict (before reading) what might be contained in the passage. This prediction was based on prior knowledge of concepts in the title. The leader also generated questions as reading proceeded and summarized the content of the passage. The leader made clarifications both during and after reading. When the student leader experienced difficulties, other students provided assistance. The leader also solicited requests for clarification from other members of the reading group. All participants were aware that they could asked for clarifications whenever they were needed. The peer leader also requested fellow students to generate questions, to summarize the segment, and to make predictions about subsequent segments of the passage.

(c) For days 7 through 11, the passages were divided into segments, with students reading each segment silently. Subjects read stories one segment at a time. Then the adult elicited oral clarifications, questions, summaries, and predictions from the students. Supportive prompting was provided as necessary; feedback was given consistently. In general, the students found this procedure engaging; they enjoyed reading the stories.

(d) For the last two days, subjects read the entire training story silently without interruption, encouraged by the adult instructor to pause occasionally to clarify, question, summarize, and predict.

Throughout instruction, the adult provided praise and feedback as well as modeled strategies when students experienced difficulties. Students were reminded repeatedly to use the trained strategies whenever they read. Although no formal measures were taken, students seemed comfortable in making the transition from external control and overt use of strategies to internal control and covert execution of the procedures. Had difficulties been apparent (and we emphasize, they were not), teaching would have been adjusted to deal with them. In short, the pace of release from external control of strategy use to internal self-regulation of the strategies seemed appropriate in that students did not manifest obvious problems.

At the end of each training session, students were given a 200-word daily assessment passage to read silently (individually administered with unlimited time to read and re-read until students felt they knew the passage). On every other day each subject was asked to retell the 200-word story aloud. On remaining days, each subject answered 10 comprehension questions aloud. Each participant was provided feedback daily. The student was shown a graph comparing the previous day's performance to all of his or her previous daily assessments.

Control condition. Control subjects also met in small groups for 13 sessions, but received no strategy training. They read the training passages silently as the instructor offered assistance with decoding and understanding of passage vocabulary when requested. This permitted a lot of interaction with the experimenter and exposure to the same materials

that reciprocally-trained students experienced. Control students were administered the daily assessments as well, with next-day feedback as in the trained condition. Thus, these subjects also had a lot of exposure to questions that could be asked about the types of passages they were reading and had practiced answering comprehension questions.

Summary

Because this was a constructive replication (Lykken, 1968) of Palincsar and Brown's (1984) Experiment 1, it is important to be clear about how the procedures for the two studies compared. Thus, important differences between the two studies are summarized in Table 3.

Insert Table 3 About Here

Results

Four dependent variables were collected. Two daily assessments were taken as training proceeded; the two standardized measures were collected both before and after the 13 training or control sessions. Each of these four variables was analyzed using a 2 (grade) by 2 (condition) by 2 (time of testing) analysis of variance with repeated measurement for the time of testing factor (Kirk, 1982). We realized from the start of the study that the effects associated with grade would be uninterpretable, however---All testing materials were different at the grade-4 and grade-7 levels (i. e., different grade-appropriate daily assessment passages & different standardized reading measures). Thus, the effects of grade were included in the analyses of variance only to remove variability associated with age and/or grade-appropriate materials. By collapsing over grade in the summary that follows, the

presentation is simplified, and thus, easier to comprehend, with absolutely no loss of information that is critical or could alter the interpretation of the significant effects that were obtained.[1] Thus, only the data for the unambiguously interpretable effects (i. e., those not involving grade) are summarized in Tables 4 (means) and 5 (test

---

Insert Tables 4 & 5 About Here

---

statistics).[2,3] For the most critical analyses in this study, the evaluations of training effects on standardized comprehension performance, the data in Tables 4 and 5 are complemented by presentation of within-grade means and statistical tests at the appropriate point in the text. (Because of the pattern of outcomes that was obtained, providing comparable means and comparisons for the other dependent variables would have been little more informative than what is already summarized in the tables, and not informative at all in ways that would have altered the interpretation of the outcomes.)

Daily Assessments

At the conclusion of each of the 13 sessions, participants read a 200-word passage. Retelling of the passage was required on half the days; for the other half of the sessions, subjects were required to answer 10 short-answer questions over the content of the passage.

Retelling. A retelling score was calculated by scoring the percentage of propositions recalled (e. g., Kintsch, 1974) from each passage. It was possible to score these retelling protocols with greater than 90% agreement for two raters. The mean percentages recalled for the three passages read and retold during the first half of training are recorded in

Table 4 as are the corresponding data for the three passages read and retold during the second half of training. The only significant effect in the analysis was that reciprocally-taught subjects recalled more than control subjects, with the effect apparent throughout training (i. e., both during the first and second halves of training).

Answering comprehension questions. On days when subjects did not retell the daily assessment passage, they responded to 10 short-answer questions over the content of the passage. These answers were scored either as correct or incorrect with no partial credit. Reciprocally-trained students outperformed controls on this variable. Overall second-half performance was better than performance during the first half of training, although, as for daily retelling, the effect of training was apparent throughout the 13 days of the experiment.

Standardized Reading Measures

In general, the pattern of outcomes on the standardized reading measures was exactly as had been hypothesized.

Comprehension. As expected, performance improved from before to after the experiment. Much more importantly, however, performance improved more in the reciprocally-trained than in the control condition. The mean pretest-to-posttest gain of 9.97 percentile points in the reciprocal instruction condition was clearly significant, t (68) = 5.14, $p < .001$; the corresponding difference of 1.63 percentile points in the control condition did not even approach statistical significance, t (68) = 0.84, $p > .40$. The same conclusion follows within each of the two grade levels: The pretest to posttest improvement in the reciprocally-trained condition was from 27.00 to 35.56 (all means in percentiles for these analyses) at the grade-4 level and from 22.67 to 34.06 at the grade-7

level, both shifts clearly significant, respective t (68) values were 3.12 and 4.15, both ps < .01. In contrast, neither pretest-to-posttest change approached significance in the control condition, larger t = 1.36, p > .10. At the grade-4 level, the mean pretest score was 28.78 compared to 28.33 at posttest; at the grade-7 level, the corresponding means were 21.17 and 24.89.

Examination of the distribution of gain scores for the comprehension measure (see Figure 1) made obvious, however, that reciprocal instruction

---

Insert Figure 1 About Here

---

was not a panacea for poor comprehenders. As suggested by Figure 1, reciprocally-trained subjects were more likely to be at or above the median (+2 percentile points) of the pretest-to-posttest difference distribution than were control students, χ^2 (1, 72) = 4.53, p < .05 (Kirk, 1982). More striking, however, a minority of reciprocally-trained children experienced large gains (e. g., only one-third of the reciprocally-trained children experienced gains of 16 percentile points or more).[4]

Vocabulary. Since 13 sessions of reciprocal instruction were not expected to affect long-term knowledge of anything except the strategies and how to use them, no significant effects on vocabulary knowledge were expected. None were obtained.

Discussion

In general, the results reported here were consistent with Palincsar and Brown's claim that reciprocal instruction improves standardized reading comprehension in students who decode adequately but do not read with high comprehension. Improvements due to reciprocal instruction

were also apparent in the daily assessment data.

We believe that this study provided a convincing demonstration of the utility of reciprocal instruction for poor comprehenders during the latter elementary-school years. The standardized comprehension measures employed here have face validity. The passages on these tests represent the types of narrative and expository contents that able students in grades 4 through 7 are expected to be able to read with high comprehension. The questions accompanying each standardized test passage tapped important skills that typify competent reading, like understanding main ideas, integration across parts of a passage, and making inferences.

Moreover, it would be very difficult to attribute the pretraining to post-training standardized comprehension gain to anything except reciprocal training, given the overall pattern of standardized test results and the way test data were collected: (a) The experiment included a control condition that received the same number of sessions with the experimenter as reciprocally-trained participants received. Control subjects also read the same number of practice passages as reciprocally-trained students and had as much practice taking tests as those in the reciprocal-training condition (i. e., they had been pretested and had received the daily assessments). That standardized comprehension improved very little for these control subjects from before to after the experiment compared to the reciprocally-trained students rules out the possibility that reciprocal-training effects were due to exposure to materials or a special teacher rather than the processes trained as part of reciprocal instruction. (b) Even though standardized comprehension improved in reciprocally-trained students, standardized vocabulary performance did not. General factors (like social influences of the

experimenter) would be expected to operate more globally. Notably, Palincsar and Brown's (1984, Experiment 1) reciprocally-trained students evidenced the same pattern of pretraining to posttraining standardized gains obtained here: There was a significant gain for comprehension from the pre-training to post-training, but very little change in standardized vocabulary performance. (c) Subjects in both conditions were given the standardized assessments simultaneously in a group. It would be difficult for experimenter demand to operate in such a situation.

Nonetheless, the overall effect on comprehension obtained here was not "breathtaking" (Carver, 1987) or "dramatic" (Pearson, 1985). The condition by time of standardized test interaction was moderate in size (i. e., 0.71 SDs; Cohen, 1977), equal to an average 6-month change in the approximate grade-equivalent score at each grade level. Nonetheless, this overall effect is quite comparable to effect sizes typically observed in metacognitive training studies (see Haller et al.'s, 1988, meta-analysis). Whether the result was dramatic for an individual varied greatly from person to person. Only one-third of the reciprocally-trained students gained more than 15 percentile points as a function of the intervention. Based on the comparison with the control condition, about half of these would have shown more than 15 percentile point gains even without reciprocal instruction. In general, the variability in gains due to reciprocal training is consistent with Palincsar and Brown (1984). Their six reciprocally-trained subjects included one who had no gain, one who had very great gain, and four in between. Of course, one obvious direction for future research would be to determine what types of children are most likely to be the beneficiaries of strategy-training packages like the one studied here.

There is no doubt that there was plenty of room for additional improvement in comprehension following reciprocal training. Perhaps there would have been greater improvement with a longer treatment (e. g., several months or a school year in duration). Evaluation of longer treatments could easily be justified since school-based instructional programs can be implemented over a very long period of time. Nonetheless, an alternative possibility is that more "bang" might be obtained for the strategy intervention "buck" with a revised version of the reciprocal instruction intervention.

Palincsar designed reciprocal instruction as the focus of her dissertation in the early 1980s, based on what was known then about how to facilitate reading comprehension. There has been an enormous amount of research on reading comprehension strategies since that initial conceptualization. There have been extremely analytical investigations in this decade that were targetted at processes encouraged by reciprocal instruction. These include studies of prior knowledge activation (e. g., Hasselhorn & Körkel, 1986), summarization (e. g., Taylor & Beach, 1984), and question generation (Davey & McBride, 1986). Revision of reciprocal instruction in light of insights from these studies might improve the treatment. In addition, evidence in favor of strategies not included in reciprocal instruction continues to mount. For instance, training poor readers to construct images representing the meaning of content that they read improves both memory (see Pressley, 1977, for a review) and comprehension (e. g., Gambrell & Bales, 1986). Teaching poor comprehenders to attend to story-grammar elements (i. e., information about story setting, characters, problems encountered by characters, attempted solutions to problems, & problem resolutions) as they read

fiction improves recall of stories (Short & Ryan, 1984). Whether reciprocal instruction would be more effective if either of these procedures (or both of them) were added to the strategy mix is an empirical question, but one that seems worth considering in light of gains produced by these strategic tactics. See Pressley, Johnson, Symons, McGoldrick, and Kurita (in press) and Lysynchuk, Pressley, d'Ailly, Smith, and Cake (1989) for comprehensive overviews of the many studies of specific strategies published in the 1980s. (In a personal communication, Palincsar indicated to us that the reciprocal intervention is now being revised to include some of the more effective strategies identified during the 1980s.)

Suffice to emphasize in closing that the outcome reported here and the Palincsar and Brown data support the conclusion that patient teaching of reading strategies can improve reading performance. When powerful strategies are taught by teachers who provide instruction that is adjusted to student difficulties, there is plenty of reason to be optimistic that measurable gains in reading competence will follow. Additional analyses of such complex reading comprehension interventions is certainly worthwhile given their obvious relevance to reading instruction.

References

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading. In P. D. Pearson (Ed.), Handbook of reading research (pp.255-291). New York NY: Longman.

August, D. L., Flavell, J. H., & Clift, R. (1984). Comparison of comprehension monitoring of skilled and less skilled readers. Reading Research Quarterly, 20, 39-53.

Bailey, R., & Quick, E. J. (1961). Young Canada readers. Toronto: Nelson.

Baker, L. (1985). How do we know when we don't understand? Standards for evaluating text comprehension. In D. L. Forrest-Pressley, G. E. MacKinnon, & T. G. Waller (Eds.), Metacognition, cognition, and human performance, Vol. 1, Theoretical perspectives (pp. 155-205). Orlando FL: Academic Press.

Bereiter, C. & Bird, M. (1985). Use of thinking aloud in identification and teaching of reading comprehension strategies. Cognition and Instruction, 2, 91-130.

Brown, A. L., & Day, J. D. (1983). Macrorules for summarizing texts: The development of expertise. Journal of Verbal Learning and Verbal Behavior, 22, 1-14.

Campbell, D. T., & Stanley, J. C. (1966). Experimental and quasi-experimental designs for research. Chicago, IL: Rand McNally.

Carver, R. P. (1987). Should reading comprehension skills be taught. In J. E. Readence, R. S. Baldwin, J. P. Konopak, & H. Newton (Eds.), Research literacy: Merging perspectives: Thirty-sixth yearbook of the National Reading Conference (pp. 115-126). Rochester NY: National Reading Conference.

Code, L. A. (1971). Individualized reading skills program. Chicago: Science Research Associates.

Cohen, J. (1977). Statistical power analysis for the behavioral sciences, 2nd edition. New York: Academic Press.

Cole, M., & Means, B. (1981). Comparative studies of how people think: An introduction. Cambridge MA: Harvard University Press.

Davey, B., & McBride, S. (1986). Effects of question-generation training on reading comprehension. Journal of Educational Psychology, 78, 256-262.

Durkin, D. (1979). What classroom observations reveal about reading comprehension instruction. Reading Research Quarterly, 14, 481-538.

Gambrell, L. B., & Bales, R. J. (1986). Mental imagery and the comprehension-monitoring performance of fourth- and fifth-grade poor readers. Reading Research Quarterly, 21, 454-464.

Haller, E. P., Child, D. A., & Walberg, H. J. (1988). Can comprehension be taught?: A quantitative synthesis of "metacognitive" studies. Educational Researcher, 17 (9), 5-8.

Hasselhorn, M., & Körkel, J. (1986). Metacognitive versus traditional reading instructions: The mediating role of domain-specific knowledge on children's text processing. Human Learning, 5, 75-90.

King, E. M. (1967). Canadian tests of basic skills. Toronto: Nelson.

Kintsch, W. (1974). The representation of meaning in memory. Hillsdale NJ: Erlbaum & Associates.

Kirk, R. E. (1982). Experimental design, 2nd Edition. Belmont CA: Brooks/Cole.

Lykken, D. T. (1968). Statistical significance in psychological research. Psychological Bulletin, 70, 151-159.

Lysynchuk, L. M., Pressley, M., d'Ailly, H., Smith, M., & Cake, H. (1988, November). A methodological analysis of reading comprehension strategy instruction research. Manuscript submitted for publication consideration. London Ontario: University of Western Ontario, Department of Psychology.

MacGinitie, W. H. (1978). Gates-MacGinitie reading test: Canadian edition. Toronto: Nelson.

Manzo, A. V. (1969). The re-quest procedure. Journal of Reading, 11, 123-126.

Meyer, B. J. F., Brandt, D. M., & Bluth, G. J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. Reading Research Quarterly, 16, 72-103.

Ninio, A., & Bruner, J. S. (1978). The achievement and antecedents of labelling. Journal of Child Language, 5, 1-15.

Palincsar, A. S. (1986). The role of dialogue in providing scaffolded instruction. Educational Psychologist, 21, 73-98.

Palincsar, A. S. (1987, April). An apprenticeship approach to the instruction of comprehension skills. Presented at the annual meeting of the American Educational Research Association, Washington DC.

Palincsar, A. M., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and monitoring activities. Cognitive and Instruction, 1, 117-175.

Palincsar, A. S., Brown, A. L., & Martin, S. M. (1987). Peer interaction in reading comprehension instruction. Educational Psychologist, 22, 231-253.

Pearson, P. D. (1985). Changing the face of reading comprehension instruction. The Reading Teacher, 38, 724-738.

Pearson, P. D., & Dole, J. A. (1987). Explicit comprehension instruction: A review of research and a new conceptualization of instruction. The Elementary School Journal, 88, 151-166.

Prescott, G. A., Balow, I. H., Hogan, T. P., & Farr, R. C. (1978). Metropolitan achievement test. New York: Harcourt Brace Jovanovich.

Pressley, M. (1977). Imagery and children's learning: Putting the picture in developmental perspective. Review of Educational Research, 47, 585-622.

Pressley, M., Johnson, C. J., Symons, S., McGoldrick, J. A., & Kurita, J. (In press). Strategies that improve memory and comprehension of what is read. The Elementary School Journal.

Pressley, M., Snyder, B. L., & Cariglia-Bull, T. (1987). How can good strategy use be taught to children?: Evaluation of six alternative approaches. In S. J. Cormier & J. Hagman (Ed.), Transfer of learning: Contemporary research and applications (pp. 81-120). Orlando FL: Academic Press.

Rinehart, S. D., Stahl, S. A., & Erickson, L. G. (1986). Some effects of summarization training on reading and studying. Reading Research Quarterly, 21, 422-438.

Rohwer, W. D., Jr. (1976). An introduction to research on individual and developmental differences in learning. In W. K. Estes (Ed.), Handbook of learning and cognitive processes, Vol. 3 (pp. 71-102). New York: Wiley.

Short, E. J., & Ryan, E. B. (1984). Metacognitive differences between skilled and less skilled readers: Remediating deficits through story grammar and attribution training. Journal of Educational Psychology, 76, 225-235.

Spache, G. D. (1972). Diagnostic reading scales. San-Francisco: McGraw-Hill.

Taylor, B. M., & Beach, R. W. (1984). The effects of text structure instruction on middle-grade students' comprehension and production of expository text. Reading Research Quarterly, 19, 134-146.

Thorn, E. A., & Richmond, M. I. (1972). Comprehension strategies-I. New York: Gage.

Vygotsky, L. S. (1978). Mind in society: the development of higher psychological processes. Cambridge, Mass. : Harvard University Press.

Wood, D. J., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. Journal of Child Psychology and Psychiatry, 17, 89-100.

Footnotes

[1] For those interested in grade effects, there were significant effects ($p < .05$) favoring grade 4s for the daily retelling, daily questions, and standard vocabulary. Based on the pretest standardized data, it was clear that the grade 4s were slightly more able than the grade 7s. This may have accounted for the advantage of grade 4s on daily retelling and question answering, or it may have been that there were slightly easier stories at the grade-4 compared to the grade-7 level. The only significant interaction involving grade was a grade by condition interaction for the daily question-answering variable. The difference between trained and control subjects was larger for grade-4 than grade-7 subjects for this variable, a finding of no obvious significance.

[2] A preliminary analysis indicated no effect for experimenter, and thus, this factor was not included in the analyses.

[3] Use of the subject as the unit of analysis (rather than the treatment reading group) is defensible because none of the assessments were gathered from the reading group as an aggregate. Even though daily assessments were obtained when the reading group met, the child was assessed individually with other group members out of sight and hearing range. All participants in the school were given the standardized reading tests simultaneously.

[4] Assuming that there was no true "gain" from pretest to posttest, some subjects would go up and some would go down due to the less than perfect reliability of the reading comprehension tests. This undoubtedly accounts

for some of the negative gain scores depicted in Figure 2.

Table 1

Summary of Tasks and Measurements as a Function of When They Occurred

---

Before the Experiment

Decoding skill measured by Diagnostic Reading Scales

Comprehension measured by Metropolitan Achievement Test for Grade 4s and Gates-MacGinitie for Grade 7s

Vocabulary subtests of the Canadian Test of Basic Skills for Grade 4s and of the Gates-MacGinitie for Grade 7s

During the Experiment

Grade-appropriate passages used to practice strategies in the reciprocal training condition; same passages reading in the control condition.

Daily assessment of comprehension and ability to retell passages on 200-word, grade-appropriate passages.

After the Experiment

Comprehension measured by Metropolitan Achievement Test for Grade 4s and Gates-MacGinitie for Grade 7s

Vocabulary subtests if the Canadian Test of Basic Skills for Grade 4s and the Gates-MacGinitie for Grade 7s

Table 2

Overview of the Four Strategies Taught As Part of Reciprocal Instruction

	Summarization
Description	A summary is one or two sentences that tells the most important ideas. A good summary does not include details or information that is not important.
Why Use It?	It can aid understanding and memory of what is read.
Some Ways to Do It?	Students told to think of what the paragraph is mostly about and to construct a sentence that reflects the most important information in the paragraph.
	Question Generation
Description	Questions are constructed about important information in the text, rather than about unimportant details.
Why Use It?	Reader can self-test to determine whether they really understand the text. By asking questions, must identify what is important in a story.
Some Ways to Do It?	Students told to select important information from text and use the words who, where, when, why, what, and how to make up questions.

Table 2 (continued)

Overview of the Four Strategies Taught As Part of Reciprocal Instruction

	Prediction
Description	Predicting involves finding clues in the structure and content of the story that might suggest what would happen next.
Why Use It?	Reader thinks about what is already known and compares it with what has already been learned in the story. Motivates reading the story to determine if predictions are correct.
Some Ways to Do It?	Students were told to use the title to make initial predictions and use clues in the story to make predictions as reading proceeds.
	Clarifying
Description	Discerning when there is a breakdown in comprehension and taking steps to restore meaning.
Why Use It?	So that the story will make sense.
Some Ways to Do It?	Students were instructed to be alert to occasions when they were not understanding the meaning of text and to process text additionally when there were problems. For instance, if a word did not make sense, students were told to try to discover the meaning of the word by reading sentences before and after it. Sometimes "or" signals the meaning of an unfamiliar word. Students were instructed to be certain they knew what referents such as "them," "it," and "they" referred to. If after rereading, something was still not clear, students were instructed to ask for assistance.

Table 3

Differences in Methods Between Palincsar and Brown (1984, Experiment 1) and the Study Reported Here

	Palincsar & Brown (1984)	This Study
Participants	Grade 7: 6 reciprocal 6 control	Grade 7: 18 reciprocal 18 control Grade 4: 18 reciprocal 18 control
	Adequate decoders, poor comprehenders. None classified as learning disabled or mentally retarded.	Adequate decoders, poor comprehenders. None classified as learning disabled or mentally retarded.
Standardized Assessments	Grade 7 comprehension: Gates-MacGinitie	Grade 7 comprehension: Gates-MacGinitie Grade 4 comprehension: Metropolitan Achievement
	Grade 7 vocabulary: Gates MacGinitie	Grade 7 vocabulary: Gates MacGinitie Grade 4 vocabulary: Canadian Tests of Basic Skills
	Group administered at pretraining to all subjects in study. Individually administered at post-training to reciprocally-trained students only.	Group administered at pretraining and post-training to all subjects in study.

Table 3 (continued)

Differences in Methods Between Palincsar and Brown (1984, Experiment 1) and the Study Reported Here

	Palincsar & Brown (1984)	This Study
Training Passages	1500 words each	300 to 900 words each
Daily Assessment Passages	400 to 475 words each	200 words each
Training	20 sessions	13 sessions
	Trained in pairs.	Trained in groups of 2 to 5 children.

Summary of Mean Values for Dependent Variables As a Function of Condition and Time of Testing

Variable	Reciprocal	Control
Daily Retelling[a]		
First Half	23.62	17.67
Second Half	22.58	15.88
Daily Questions[b]		
First Half	50.09	41.32
Second Half	61.36	45.30
Standardized Comprehension[c]		
Before	24.84	24.97
After	34.81	26.61
Standardized Vocabulary[c]		
Before	26.16	25.06
After	31.13	31.17

[a] Mean percentage of passage propositions that were recalled.

[b] Mean percentage of questions answered correctly.

[c] Mean percentile score.

F Values for Analyses of Dependent Variables (Interpretable Effects Only)

Effect Tested	Daily Retelling	Daily Questions	Standard Comprehension	Standard Vocabulary
Condition (C)	4.68*	14.00***	1.45	0.02
Time of Testing (T)				
1st Vs. 2nd Half	2.52	12.43***	----	----
Before Vs. After	----	----	8.95**	3.54
C X T	0.18	2.83	4.61*	0.04
MS_E for C effect (Between Subjects)	302.91	390.48	402.79	656.70
MS_E for T and C X T effects (Within Subjects)	27.95	165.96	135.61	307.75

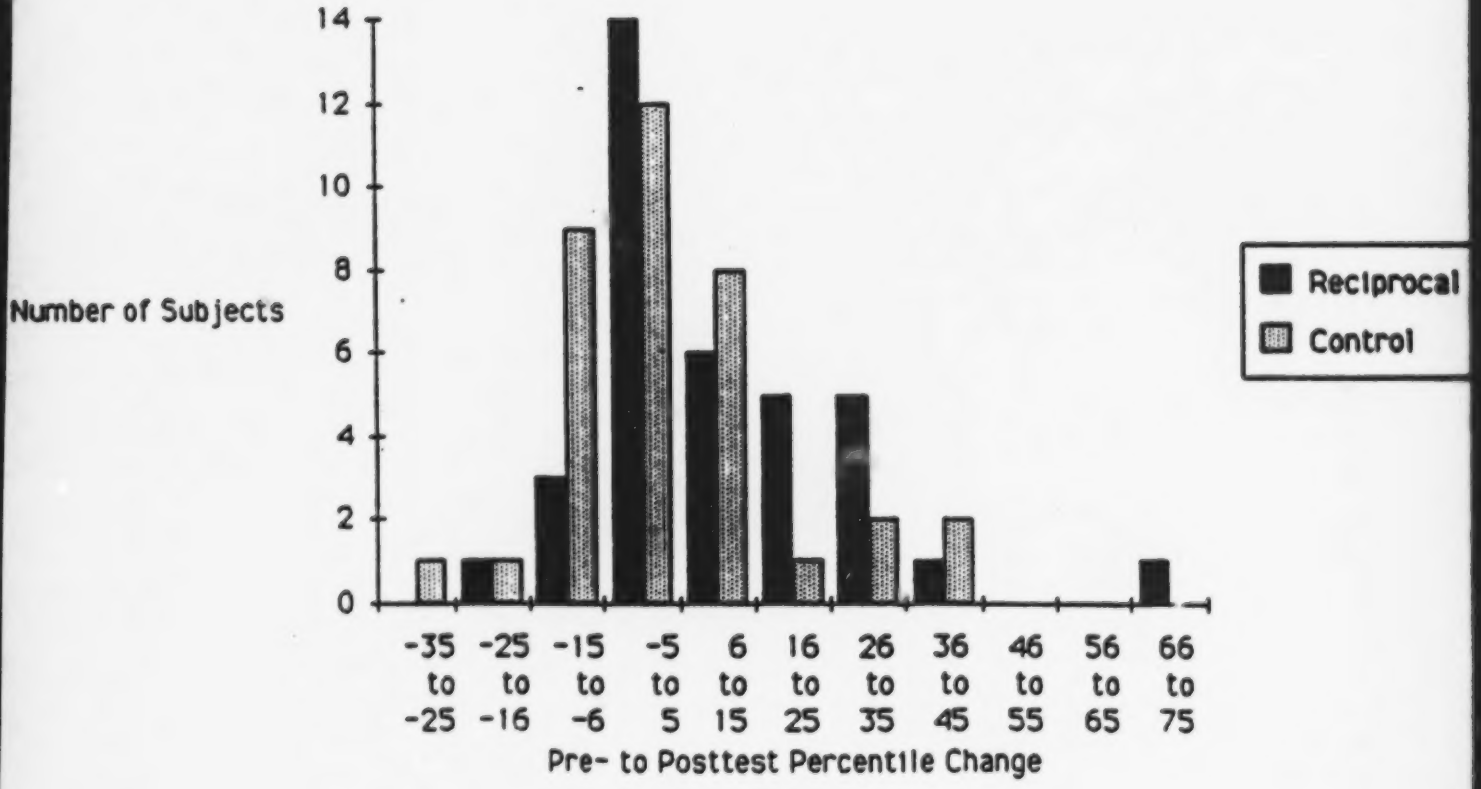
Note. Numerator degrees of freedom = 1; denominator degrees of freedom = 68.

* $p < .05$

** $p < .01$

*** $p < .001$

Figure 1: Distribution of Pretest to Posttest Percentile Changes on Standardized Reading Comprehension



Appendix A

Sample Daily Comprehension Passage

The violence of an earthquake depends upon the nature of the rock movement and on its location. Very small shifts, whether near the earth's surface or deeper in the crust, cause only a tremor. Larger shifts that are far below will have a slight effect. But a large, shallow movement will have powerful effects.

Earthquakes are most likely to occur where the earth's surface is most uneven. One such area exists where high mountains are located close to very low areas. The low areas may be valleys on land or nearby deeps in the ocean. Both the mountains and the deeps may be under the surface of the ocean.

Regions of very uneven surface are called earthquake belts. One large earthquake belt lies below the Pacific coasts of North and South America and long the coasts of Japan and other islands of the Pacific. Other earthquake regions are in the East Indies, the West Indies, Portugal, Spain, southern Italy, Greece, and Turkey and long the southern edge of the Himalayan Mountains. People in earthquake belts regularly feel small tremors and at times are subjected to violent quakes (Code, 1971, pp. 141-142).

Comprehension Questions

(1) On what does the violence of an earthquake depend?

(2) What are tremors caused by?

(3) Where are earthquakes most likely to occur?

(4) What are "earthquake belts"?

(5) Where may earthquakes be found?

(6) Why would one not want to live in Japan?

(7) What is the story about?

(8) Where is the earth's surface uneven?

(9) What causes the worst earthquakes?

(10) What is the earth like in Spain?